UNITÉ DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél. (1) 39 63 55 11

Rapports de Recherche

N° 1339

Programme 5
Automatique, Productique,
Traitement du Signal et des Données

FAST SOLUTION OF GENERAL NONLINEAR FIXED POINT PROBLEMS

Roberto L. V. GONZALEZ
Mabel M. TIDBALL

Décembre 1990

RR-1339

FAST SOLUTION OF GENERAL NONLINEAR FIXED POINT PROBLEMS

SOLUTION RAPIDE DES PROBLEMES DE POINT FIXE NON LINEAIRES

Roberto L. V. González * - Mabel M. Tidball *+

* Instituto de Matemática Beppo Levi
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario
Avda. Pellegrini 250 – 2000 Rosario – Argentine.

+ This paper is included in the activities developed in the frame of the Cooperation Project INRIA-Instituto de Matemática Beppo Levi
Coordinators of the project: E. Rofman - R. González

ABSTRACT

In this paper, we develope a general procedure to stabilize the usual Newton method in such a way that algorithms obtained always converge to the unique solution of the problem.

In general, although the modified Newton's algorithm is convergent, no improvement of the order of speed of convergence can be expected (compered to usual iterative algorithm with geometric rate of convergence); in fact we give an example where, independently of the chosen starting point, the convergence is geometric of order 1/3.

In spite of these negative results, two fields of successful application are shown: the case where the operator $T \in C^1 \bigcap H^{2,\infty}$ and the case where T is polyhedric. In the first case, quadratic convergence is proved; in the second one convergence in a finite number of steps is obtained.

Finally, numerical results are shown for an example issued from the field of differential games.

RESUME

On présente ici un procédé general pour la stabilisation de l'algorithme de Newton, la modification a été faite de telle façon que l'algorithme obtenu converge de n'importe quel point, à l'unique solution du problème originel.

Généralement, malgré que l'algorithme de Newton modifié soit toujours convergent, aucune amélioration de la vitesse de convergence peut être espéré (par comparaison avec les algorithmes usuels dont la convergence est géométrique); pour démontrer célà nous présentons un contre-example où la convergence est géométrique d'ordre 1/3 indépéndantement du point initial choisi.

Malgré ces résultats décourageants, deux domaines d'application réussis sont présentés: le cas où l'opérateur $T \in C^1 \bigcap H^{2,\infty}$ et le cas où T est polyédrique. Pour le prémier cas la convergence est quadratique et dans le deuxième on a convergence dans un nombre fini d'itérations.

Finalement, on présente des résultats numériques pour un problème issue du domaine de jeux différentiels.

CONTENTS

Lasciate ogni speranza, voi ch'entrate

Dante Alighieri

La Divina Commedia

1. INTRODUCTION

Frequently, optimal control problems and differential games problems originate variational inequalities (see [7] and [13]). Also, problems issued from other fields are reduced to these type of inequalities. In order to obtain numerical solutions, it is necessary to discretize the original problem (the continuous solution is contained in an infinite dimensional space); in that way, the final problem, which must be solved computationally, is reduced to find the fixed point of a contractive operator. When the actualization rate of the original problem is small (see [8]), the numerical resolution (found by relaxation type iterative algorithms, see [1]) may lead to slowly convergent procedures. In [8], [9], [10], we have introduced acceleration procedures to improve the speed of convergence of the usual algorithm of Picard type; it essentially consists in the combination of Picard's and Newton's methods. In this paper, we extend the procedures presented there, in order to make them applicable to general nonlinear contractive operators (where they no longer are the consecuence of the discretization of differential games or optimal control problems).

The set of results obtained is the following. In a first place we have developed a general procedure to stabilize the usual Newton method in such a way that algorithms obtained always converge to the unique solution of the discrete problem (in particular, this technique enables us to transform Howard's methods, which are not convergent in the case of general differential games problems, and to make them applicable to others problems outside the original fields of application). In general, although the modified Newton's algorithm is convergent, no improvement of the order of speed of convergence can be expected; in fact we give an example where independently of the chosen starting point, the convergence is geometric of order $1/3$.
In spite of these negative results, two fields of successful application are shown: the case where the operator $T \in C^1 \bigcap H^{2,\infty}$ and the case where T is polyhedric. In the first case, quadratic convergence in proved; in the second one convergence in a finite number of steps is obtained.

Finally, numerical results are shown for an example issued from the field of differential games.

2. PROBLEM DESCRIPTION

2.1 Elements of the Problem.

Let T be an operator defined in $\Re^n$, such that

$$T \in C^o \bigcap H^{1,\infty}(\Re^n) \tag{1}$$

We assume that operator T is contractive, i.e. there exist ρ, $0<\rho<1$ such that T verifies

$$\|Tx - T\hat{x}\| \leq (1-\rho)\|x - \hat{x}\| \quad \forall\ x, \hat{x} \in \Re^n. \tag{2}$$

The algorithms proposed in this paper, are aimed to compute in a fast way the solution of the following problem:

Problem P:

$$\boxed{\textit{Find } \overline{x} \in \Re^n, \textit{ such that } T\overline{x} = \overline{x}.} \tag{3}$$

2.2 Existence and Uniqueness of Solution.

Proposition 2.1: *There exists an unique solution for* (3).

Proof:The proof is trivial (see, for example [15])

□

2.3 Iterative Computation of the Fixed Point.

The Fixed Point Theorem gives us the following algorithm for the computation of $\bar{x}$:

A0 algorithm:

Step 1: set $x^0 \in \Re^n$, and $\nu=0$.

Step 2: compute $x^{\nu+1} = Tx^{\nu}$

Step 3: if $x^{\nu} = x^{\nu+1}$ then, stop; else, set $\nu=\nu+1$ and go to Step 2.

For the convergence of algorithm A0 the following result holds (see [1]):

Theorem 2.1: *A0 algorithm produces either a finite sequence x^{ν} whose last element is the exact solution $\bar{x}$ of the problem, or generates an infinite sequence x^{ν} converging to $\overline{x}$. Also, the following bound for the approximation error is valid:*

$$\|x^{\nu} - \bar{x}\| \leq (1-\rho)^{\nu}\|x^0 - \bar{x}\|. \tag{4}$$

3. AN ABSTRACT ALGORITHM AND ITS CONVERGENCE

3.0 Preliminary Discusion.

Although algorithm A0 converges from any arbitrary initial point x_0 , the corresponding speed of convergence is very slow when factor ρ tends to zero. To accelerate this procedure, Newton's type methods should be used. But in general, these methods are not convergent from everywhere and in consequence, it is necessary to design a technic to stabilize them and to achieve globally convergence,(see for instance the appendix, where a particular case of Newton's methods, Howard's method; originally introduced to solve optimal control problems, may be not convergent when it is applied to solve differential games problems).

To stabilize the method we use a merit function which measures the distance from the current point x to the solution $\bar{x}$. The special algorithm presented here generates a sequence of points x^{ν} such that the associated sequence $V(x^{\nu})$ is a monotonically decreasing sequence converging to zero. This procedure is obviously related to Lyapunov's methodology to stabilize dynamical systems. (see for illustrative remarks about this fact, the clever introduction of the book of Polak [16])

3.1 Lyapunov's Function. Equivalent Problem.

We define, in a natural way, the following Lyapunov's function

$$V(x) = \| Tx - x \|^2 \tag{5}$$

The function V satisfies the following properties:

$$V(x) = 0 \Leftrightarrow x = Tx \tag{6}$$

$$V(x) \geq \rho^2 \| x - \bar{x} \|^2 \tag{7}$$

where $\bar{x}$ is the solution of the problem. From (5), (6) obviously holds, also as

$$\|x-Tx\| = \|x-Tx + T\bar{x} - \bar{x}\| \geq \|x-\bar{x}\| - \|Tx-T\bar{x}\| \geq$$

$$\geq \|x-\bar{x}\| - (1-\rho)\|x-\bar{x}\| = \rho\,\|x-\bar{x}\|$$

then

$$V(x) = \| x - Tx \|^2 \geq \rho^2 \, \|x - \bar{x}\|^2$$

We are now in conditions to introduce the auxiliary

Problem P':

$$\boxed{\text{Find } \bar{x} \in \Re^n, \text{ such that } V(\bar{x}) = \min\{ V(x) : x \in \Re^n\}} \tag{8}$$

It is obvious, by (6) and (7) that problems P and P' are equivalent in the sense that both of them have the same solution $\bar{x}$.

3.2 Abstract Algorithm.

We define here a general algorithm and we prove the convergence in terms of the descent of Lyapunov's function.

Let M be a map such that:

$$M: \Re^n \rightarrow \mathcal{P}(\Re^n)$$

We shall suppose that M is a decreasing transformation of V in following sense:

$$V(y) \leq \gamma \, V(x) \;\; \forall \; y \in M x \tag{9}$$

where $0 \leq \gamma < 1$.

Algorithm Aa

Step 1: set $x^o \in \Re^n$, and $\nu = 0$.

Step 2: choose $x^{\nu+1} \in M x^{\nu}$

Step 3: if $x^{\nu} = x^{\nu+1}$, then stop; else set $\nu = \nu + 1$ and go to Step 2.

The convergence of algorithm Aa is assured by condition (9), as it is established in the following

Theorem 3.1: *If* $V(y) \leq \gamma V(x) \;\; \forall \; y \in M x$ *with* $\gamma < 1$, *then the abstract algorithm* Aa *gives the solution* $\bar{x}$ *in a finite number of steps or generates a sequence converging to* $\bar{x}$.

Proof: For any initial point x_0, if algorithm Aa finishes in a finite number of steps, we have:

$$V(x^{\bar{\nu}}) = V(x^{\bar{\nu}+1}) \leq \gamma \, V(x^{\bar{\nu}})$$

and that implies

$$V(x^{\bar{\nu}}) = 0$$

so, by (6)

$$x^{\bar{\nu}} = \bar{x}$$

If sequence x^{ν} is infinite, we have

$$V(x^{\nu}) \leq \gamma^{\nu} V(x^{0})$$

so

$$V(x^{\nu}) \to 0$$

that implies, by (7) that $\| x^{\nu} - \bar{x} \| \to 0$, because

$$\| x^{\nu} - \bar{x} \|^2 \leq \frac{1}{\rho^2} V(x^{\nu})$$

$\square$

3.3 Necessity of Condition $V(M(x)) \leq \gamma\, V(x)$.

In algorithm Aa, condition

$$V(y)) \leq \gamma V(x) \quad \forall\, y \in M(x) \tag{10}$$

cannot be replaced by

$$V(y) < V(x)$$

without losing the property of convergence.

In effect, let us consider the following function $T : \Re \to \Re$

$$T(x) = \begin{cases} -\left(\frac{3x+1}{4}\right)^{\frac{2}{3}} + x & \text{if } x > 1 \\ 0 & \text{if } -1 \leq x \leq 1 \\ \left(\frac{1-3x}{4}\right)^{\frac{2}{3}} + x & \text{if } x < -1 \end{cases}$$

Then, $\| T' \| < 1$ and problem P has the unique solution $\bar{x} = 0$.

If we define the map

$$M(x) = 0 \qquad \text{if } x \in [-1,1]$$

$$M(x) = x - (1 - T'(x))^{-1}(T(x) - x) \qquad \text{if } x \notin [-1,1]$$

we obtain

- if $x^{\nu} > 1$ then

$$M(x^{\nu}) = x^{\nu} + (1 - T'(x^{\nu}))^{-1}\,(T(x^{\nu}) - x^{\nu}) = -\left(\frac{x^{\nu}+1}{2}\right) < -1$$

- if $x^{\nu} < -1$ then

$$x^{\nu+1} = x^{\nu} + (1 - T'(x^{\nu}))^{-1} (T(x^{\nu}) - x^{\nu}) = \frac{1 - x^{\nu}}{2} > 1$$

In all cases, the following property can be proved without difficulty

$$V(M(x^{\nu})) < V(x^{\nu});$$

however, if $\|x_0\| > 1$, algorithm Aa generates a sequence $\{x^{\nu}\}_{\nu=1}^{\infty}$ such that, although sequence $\|x^{\nu}\|$ is decreasing, it is not convergent to zero (in fact, $\|x^{\nu}\| = 1 + (\|x^0\| - 1)\, 2^{-\nu}$). That sequence has two cluster points, 1 and -1, while sequence $V(x^{\nu})$ converges monotonically to 1.

3.4 Practical Algorithms.

We have presented above the general algorithm Aa that converges from everywhere. Now we shall define two practical implementation of it, algorithms A1 and A2, trying that these algorithms apply, whenever possible or convenient, Newton's method to solve the non linear equation $Tx - x = 0$.
This situation is detected testing the descent of Lyapunov's function V. When Newton's method does not produce a decrement of V, Newton's direction and direction $Tx - x$ (given by algorithm A0) are associated, until the new computed point $x^{\nu+1}$ satisfies condition $V(x^{\nu+1}) \leq \gamma V(x^{\nu})$.
In A1 algorithm, this condition is defined in Step 3, and it involves the computation of T(T(x)). Algorithm A2 avoids the computation of T(T(x)), using an adaptative estimation of factor γ.

3.4.1 *Preliminaries for the application of Newton's method.*

Definition of the set of "differentials" $\Theta(x)$.

As operator T is Lipschitz continuous, it is only almost everywhere differentiable. In order to define in a correct way algorithms A1 and A2 (introduced in the following section), it is necessary at every point of $\Re^n$ to define generalized linear operators (in fact, we use here a restricted version of Clarke's subdifferential or peridifferential of T at x, for details and a discusion about this maters see [3], [12]) such that they coincide with T'(x) at points where T is continuously differentiable. With this aim, we introduce the following concepts:

Definition 1

$$\hat{\Theta}(x) = \{T'(x)\} \quad \text{if T is differentiable in x}$$

$$\hat{\Theta}(x) = \emptyset \quad \text{if T is not differentiable in x}$$

Definition 2

$$\Theta(x) = \bigcap_{\epsilon \geq 0} \overline{\bigcup_{y} \{ \hat{\Theta}(y) \,/\, \| y - x \| \leq \epsilon \}} \tag{11}$$

By (1) and (2) we have that T is differentiable almost everywhere and that in any point where T is differentiable it is satisfied that

$$\| T' \|_{\infty} \leq 1 - \rho,$$

in consequence it is easy to prove (see [12]) the following properties:

$$\Theta(x) = T'(x) \quad \text{if T is continuously differentiable at x}$$

$$\Theta(x) \neq \emptyset \qquad \forall\, x$$

$$\forall\, \tau \in \Theta(x), \| \tau \| \leq 1 - \rho \tag{12}$$

3.4.2 *Algorithm A1. Definition and properties.*

Algorithm A1

Step 0:

Give a sequence λ_p , η_p / $\lambda_1 = 1$, $\eta_1 = 0$, $\lambda_p \to 0$, $\eta_p \to 1$ as $p \to \infty$

set $\nu = 0$, $x^\nu = x_0$

Step 1:

If $T(x^\nu) = x^\nu$; then, stop

else,

set $p=1$, $\hat{\beta}_\nu = \dfrac{V(T(x^\nu))}{V(x^\nu)}$,

choose an arbitrary $T' \in \Theta(x^\nu)$

and set $v^\nu = (I-T')^{-1} (T(x^\nu) - x^\nu)$,

$w^\nu = T(x^\nu) - x^\nu$

Step 2:

set

$v^{\nu,p} = \lambda_p v^\nu + \eta_p w^\nu$,

$y^{\nu,p} = x^\nu + v^{\nu,p}$

Step 3:

If $V(y^{\nu,p}) < \dfrac{1+\hat{\beta}_\nu}{2} V(x^\nu)$; then $x^{\nu+1} = y^{\nu,p}$, $\nu = \nu+1$ go to step 1

else $p=p+1$ and go to step 2

Remark:

We denote M(x) the set of points given by algorithm A1. As $\Theta(x)$ is not single-valued, in general, also the set of points generated by algorithm A1 is not a singleton.

Theorem 3.2: *The loop 2-3-2 always finishes in a finite number of steps, defining for each point* x^ν *a* *new point* $M(x^\nu) = x^{\nu+1}$. *The operator* M *verifies property* (9) *and then* x^ν *converges to* $\bar{x}$.

Proof: To prove that test of step 3 is satisfied, making possible the generation of a new element $x^{\nu+1}$, it must be verified that: $\forall\ \nu,\ \exists\ p\ /$

$$V(y^{\nu,p}) \leq \frac{1+\hat{\beta}_\nu}{2}\ V(x^\nu) \tag{13}$$

due to

$$y^{\nu,p} = x^\nu + \lambda_p v^\nu + \eta_p w^\nu$$

we have

$$\lim_{p\to\infty} y^{\nu,p} = T(x^\nu)$$

As T is contractive we know that

$$V(Tx^\nu) = \| T(Tx^\nu) - Tx^\nu \|^2 \leq (1-\rho)^2 \| Tx^\nu - x^\nu \|^2 = (1-\rho)^2 V(x^\nu)$$

By definition

$$V(Tx^\nu) = \hat{\beta}_\nu V(x^\nu)$$

we have

$$\hat{\beta}_\nu \leq (1-\rho)^2 < 1$$

so

$$\lim_{p\to\infty} V(y^{\nu,p}) = V(T(x^\nu)) = \hat{\beta}_\nu V(x^\nu) < \frac{1+\hat{\beta}_\nu}{2} V(x^\nu)$$

in consequence $\forall\ x^\nu,\ \exists\ p$ such that (13) is verified.

So, we have proved that the test of step 3 is satisfied, generating a new element $x^{\nu+1}$. We shall see now that this sequence converges to $\bar{x}$, solution of our problem.
As test of step 3 is verified, we know

$$V(x^{\nu+1}) \leq \frac{1+\hat{\beta}_\nu}{2} V(x^\nu) \leq \gamma V(x^\nu)$$

with

$$\gamma = \frac{1+(1-\rho)^2}{2} < 1$$

and by theorem 3.1, sequence $x^\nu \to \bar{x}$ when $\nu \to +\infty$.

$\square$

3.4.3. *Algorithm A2. Definition and properties.*

Algorithm A2

Step 0:

Give a sequence λ_p , η_p / $\lambda_1 = 1$, $\eta_1 = 0$, $\lambda_p \to 0$, $\eta_p \to 1$ when $p \to \infty$

set $\nu = 0$, $x^\nu = x_0$, $\alpha = \alpha_0 \in [0, 1)$, $\tilde{p}=1$

Step 1:

If $T(x^\nu) = x^\nu$;then, stop

else,

choose an arbitrary $T' \in \Theta(x^\nu)$, $p=1$, $\alpha_{\nu+1} = \alpha_\nu$

and set

$$v^\nu = (I-T')^{-1} (T(x^\nu) - x^\nu)$$
$$w^\nu = T(x^\nu) - x^\nu$$

Step 2:

set

$$v^{\nu,p} = \lambda_p v^\nu + \eta_p w^\nu$$
$$y^{\nu,p} = x^\nu + v^{\nu,p}$$

Step 3:

set

$$\alpha_p^{\nu+1} = \alpha_\nu \quad \text{if } p \le \tilde{p}$$
$$\alpha_p^{\nu+1} = \alpha_{\nu+1} \quad \text{if } p > \tilde{p}$$

Step 4:

If $V(y^{\nu,p}) < \alpha_p^{\nu+1} V(x^\nu)$;then, $x^{\nu+1} = y^{\nu,p}$, $\bar{p}^\nu = p$,

$\nu = \nu+1$ go to step 1

else, $p=p+1$,

if $p>\tilde{p}$; then,

$\tilde{p}=\tilde{p}+1$, $\alpha_{\nu+1} = \dfrac{1+\alpha_{\nu+1}}{2}$

and go to step 2

else, go to step 2

Theorem 3.3: *The loop 2-3-4-2 always finishes in a finite number of steps, defining for each point x^ν a new point $M(x^\nu) = x^{\nu+1}$. The operator M verifies property (9), (with $\gamma = 1 - \frac{1-\bar{\alpha}}{2^{n_0}}$) and then x^ν converges to $\bar{X}$.*

Proof: We shall prove that loop 2-3-4-2 always finishes in a finite number of steps generating a new element $x^{\nu+1}$. With this aim, we define:

$$\xi(0) = \alpha\,, \text{ with } 0 \le \alpha < 1$$

$$\xi(p) = \frac{1 + \xi(p-1)}{2}$$

then

$$\xi(p) = 1 - \frac{1-\alpha}{2^p}, \ \forall\, p \ge 1$$

By algorithm A2, we have

$$\alpha_p^\nu \le \xi(p), \quad \forall\, p \ge 1$$

by definition of $y^{\nu,p}$ and properties of operator T, we have:

$$\left| y^{\nu,p} - T(x^\nu) \right| \le \Big((1-\rho)(1-\eta_p) + \tfrac{1}{\rho}\lambda_p(2-\rho) + (1-\eta_p) \Big) | x^\nu | = \phi_p | x^\nu |$$

where

$$\phi_p = (1-\rho)(1-\eta_p) + \tfrac{1}{\rho}\lambda_p(2-\rho) + (1-\eta_p)$$

then

$$\phi_p \to 0, \text{ as } p \to \infty$$

we know also that $\forall\, z, \bar{z} \in \Re^n$

$$V(z) - V(\bar{z}) = | T(z) - z |^2 - | T(\bar{z}) - \bar{z} |^2 =$$

$$= \Big(| T(z) - z | + | T(\bar{z}) - \bar{z} | \Big) \Big(| T(z) - z | - | T(\bar{z}) - \bar{z} | \Big) \le$$

$$\le (2-\rho) \Big(|z| + |\bar{z}| \Big) | T(z) - T(\bar{z}) - z + \bar{z} | \le$$

$$\le (2-\rho)^2 \Big(|z| + |\bar{z}| \Big) | z - \bar{z} |$$

in consequence, as

$$V(T(x^\nu)) \le (1-\rho)^2 V(x^\nu)$$

we have

$$\frac{V(y^{\nu,p})}{V(x^{\nu})} \leq (1-\rho)^2 + \frac{V(y^{\nu,p}) - V(T(x^{\nu}))}{V(x^{\nu})} \leq$$

$$\leq (1-\rho)^2 + \frac{(2-\rho)^2\left(\| y^{\nu,p}\| + \| T(x^{\nu})\|\right)}{\rho^2 \| x^{\nu}\|} \| y^{\nu,p} - T(x^{\nu})\| \leq$$

$$\leq (1-\rho)^2 + \frac{(2-\rho)^2\left(2-\rho+\phi_p\right)}{\rho^2}\phi_p$$

then, there exist $n_0 \;/\; \forall\ p \geq n_0$

$$(1-\rho)^2 + \frac{(2-\rho)^2\left(2-\rho+\phi_p\right)}{\rho^2}\phi_p \leq \xi(p)$$

hence, the test of step 3 is verified $\forall\ p \geq n_0$, and as a consequence loop 2-3-4-2 finishes in not more than n_o steps, i.e. $\bar{p}^{\nu} \leq n_0$.

It is obvious that $\alpha_p^{\nu} \leq 1 - \frac{1-\alpha}{2^{n_o}}$, so

$$V(y^{\nu,\bar{p}}) = V(x^{\nu+1}) \leq \alpha_p^{\nu+1} V(x^{\nu}) \leq \left(1 - \frac{1-\alpha}{2^{n_o}}\right) V(x^{\nu})$$

Theorem 3.3 holds because hypothesis of theorem 3.1 is verified with $\gamma = 1 - \frac{1-\alpha}{2^{n_o}}$

□

3.4.4 *Necessity of the use of direction* Tx − x.

Algorithms A1 and A2 are based in the common use of directions b_1 and b_2.

$$b_1 = Tx - x$$

$$b_2 = (I - T')^{-1}(Tx - x)$$

This combination ensures the global convergence of algorithms through the descent of Lyapunov's function V. Although in the case where T is differentiable, Newton's direction b_2 is always a descent direction and the search could be restricted to that line, we shall show that in the case where T is not differentiable we cannot use only Newton's direction b_2, because it may be not a descent direction.

Newton's methods proposed in this paper are based in choosing a matrix $T' \in \Theta(x)$; if T is not differentiable, $\Theta(x)$ has more than a unique element. If we choose $T' \in \Theta(x)$ it may occur that the new direction is not a descent direction for function V, as it is shown in the counterexample given below. In order to avoid this phenomenon and to get a stable and globally covergent algorithm. Algorithms A_1 and A_2 take a suitable combination of Newton's direction and direction b_1, that always brings a descent direction.

Counterexample where w is not a descent direction

Let $T: \Re^2 \to \Re^2$ be such that

$$Tx = \begin{cases} Mx + p & \text{if } x_1 \geq 0 \\ \hat{M}x + p & \text{if } x_1 \leq 0 \end{cases}$$

where

$$\hat{M} = \begin{bmatrix} 0.9 & 0 \\ -0.9 & 0 \end{bmatrix} \qquad M = \begin{bmatrix} 0.9 & 0 \\ 0.9 & 0 \end{bmatrix}$$

$$p = \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix}$$

By definition of T, we have that in the set $\{x \in \Re^2 / x_1 = 0\}$ (common boundary of the individual domains where T is defined as an affine function) operator T is well defined and it is continuous.
It is clear that at $x=0$, $\Theta(x) = \{M, \hat{M}\}$. When we apply algorithms A1 or A2, if the element chosen by them is $\hat{M}$, we can see that direction $b_2 = (I - \hat{M})^{-1}(Tx - x)$ generates a half-line contained in the set where T is an affine function with kernel M, in effect:

$$b_2 = (I - \hat{M})^{-1}\, T(0) = (I - \hat{M})^{-1}\, p = \begin{bmatrix} 10 & 0 \\ -9 & 1 \end{bmatrix} \begin{bmatrix} 0.1 \\ 0.9 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

so, for the derivative of V in the direction b_2, we have

$$\frac{\partial V}{\partial b_2} = (\nabla V, b_2) = -p'(M - I)(\hat{M} - I)^{-1} p = -b_2'(\hat{M} - I)'(M - I)\, b_2 \tag{14}$$

In this case

$$(\hat{M}-I)'(M-I)=\begin{bmatrix}-0.8 & 0.9\\ \\ -0.9 & 1\end{bmatrix}$$

in consequence

$$\frac{\partial V}{\partial b_2}=(\nabla V, b_2)=-p'(M-I)(\hat{M}-I)^{-1}p=-b_2'(\hat{M}-I)'(M-I)b_2=0.8>0$$

and Newton's direction b_2 is not a descent direction.

□

4. SPECIAL CASES

4.1 Quadratic Convergence.

When operator T is smoother than in the general case; i.e, strictly

$$T \in C^1 \cap H^{2,\infty} \tag{15}$$

we have that algorithms A1 and A2 converge globally from any starting point with quadratic rate of convergence, i.e.

$$\| x^{\nu+1} - \bar{x} \| \leq K \| x^{\nu} - \bar{x} \|^2$$

The proof of this property is standard and we include it here only by sake of completeness.

Theorem 4.1: *If* (15) *holds; then, there exist* $K > 0$, $\hat{\nu}(x_0)$ *such that*

$$\|x^{\nu+1} - \bar{x}\| \leq K \|x^{\nu} - \bar{x}\|^2 \quad \forall\ \nu \geq \hat{\nu}(x_0)$$

Proof: We shall prove that there exists an index $\bar{\nu}$ such that for $\nu > \bar{\nu}$, both algorithms A1 and A2 produce as successive points the pairs given by Newton's rules, i.e.

$$x^{\nu+1}=x^{\nu}+(I-T'(x^{\nu}))^{-1}.(T(x^{\nu})-x^{\nu})$$

in effect, if we define

$$Q(x) = x + (I - T'(x))^{-1}.(T(x) - x)$$

we have :

$$\| Q(x) - \bar{x} \| \leq K \| x - \bar{x} \|^2 \tag{16}$$

because

$$T(x) = \bar{x} + T'(\bar{x}).(x - \bar{x}) + \phi(x)$$

with

$$\| \phi(x) \| \leq K \| x - \bar{x} \|^2$$

Also

$$T'(x) = T'(\bar{x}) + W(x)$$

with

$$\| W(x) \| \leq K \| x - \bar{x} \|$$

In consequence, we obtain

$$T(x) = \bar{x} + T'(x).(x - \bar{x}) + \psi(x)$$

where

$$\| \psi(x) \| \leq K \| x - \bar{x} \|^2$$

then, finally we have

$$Q(x) = x + (I - T'(x))^{-1}.(Tx - x) = x + (I - T'(x))^{-1}\Big((\bar{x} - x + T'(x).(x - \bar{x}) + \psi(x)\Big) =$$

$$= x + (I - T'(x))^{-1}\Big((T'(x) - I).(x - \bar{x}) + \psi(x)\Big) = \bar{x} + (I - T'(x))^{-1}\psi(x)$$

which implies (16).

In consequence

$$\| T(Q(x^{\nu})) - T(\bar{x}) \| \leq (1 - \rho) \| Q(x^{\nu}) - \bar{x} \| \leq (1 - \rho) K \| x^{\nu} - \bar{x} \|^2 \tag{17}$$

we take now into account the following relation

$$\| T(Q(x^\nu)) - Q(x^\nu) \| \leq \| T(Q(x^\nu)) - T(x) \| + \| T(x) - Q(x^\nu) \|$$

By (16)

$$\| T(\bar{x}) - Q(x^\nu) \| = \| \bar{x} - Q(x^\nu) \| \leq K \| x^\nu - \bar{x} \|^2, \tag{18}$$

and by (17)

$$\| T(Q(x^\nu)) - T\bar{x} \| \leq (2-\rho) K \| \bar{x} - x^\nu \|^2,$$

in consequence, by virtue of (7)

$$V(Q((x^\nu))) = \| T(Q(x^\nu)) - T(\bar{x}) \|^2 \leq (2-\rho)^2 K^2 \| \bar{x} - x^\nu \|^4$$

$$\leq \frac{(2-\rho)^2 K^2 \| \bar{x} - x^\nu \|^2}{\rho^2} V(x^\nu) < \alpha_1^{\nu+1} V(x^\nu)$$

when the following condition holds:

$$\frac{(2-\rho)^2 K^2 \| \bar{x} - x^\nu \|^2}{\rho^2} < \alpha_1^{\nu+1} \tag{19}$$

As $x^\nu \to \bar{x}$, and $\alpha_1^{\nu+1}$ is uniformly bounded from below, $\exists\ \bar{\nu}\ /\ \forall\ \nu \geq \bar{\nu}$ (19) is verified and in consequence, $\forall\ \nu \geq \bar{\nu}$ the test of step 3 of algorithm A1 and the first test of step 4 of algorithm A2 are satisfied and we have the following equality

$$x^{\nu+1} = Q(x^\nu)$$

Then by (18), we obtain

$$\| x^{\nu+1} - \bar{x} \| \leq K \| x^\nu - \bar{x} \|^2$$

and so a quadratic rate of convergence holds.

$\square$

4.2 Convergence in a Finite Number of Steps.

In many problems, for example those originated in discretization of differential games or of equations or non linear inequalities, operator T results locally affine, i.e. continuous and affine is its restriction to some determined sets. In this case we call T a polyhedrical operator; strictly, we define T as polyhedrical if there exist a finite set (with cardinality χ) of indices "q" and for each q there is a vector $a_q \in \Re^n$, a nxn matrix M_q and a set $S_q \subset \Re^n$ such that the following properties hold:

$$\bigcup_{q=1}^{\chi} S_q = \Re^n \tag{20}$$

$$M_q x + a_q = M_{q'} x + a_{q'} \;\; \forall \, x \;/\; x \in S_q \bigcap S_{q'} \tag{21}$$

So, it follows that T is a well defined and continuous operator such that:

$$Tx = M_q x + a_q \quad \forall \; x \in S_q$$

Properties:

$$\Theta(x) \subset \{M_q / q \in \hat{Q}(x)\}$$
$$\hat{Q}(x) = \{q / Tx = M_q x + a_q\}$$

Proof: The proof is trivial and we shall omit it.

Theorem 4.2: *If* T *is polyhedrical, algorithms* A1 *and* A2 *converge in a finite number of steps.*

Proof: Let $\bar{x}$ be / $T\bar{x} = \bar{x}$, we define :

$$\bar{Q} = \{q \;/\; T\bar{x} = M_q \, \bar{x} + a_q \}$$

We call $\hat{q}(\nu)$ the index such that the element $T'_\nu \in \Theta(x^\nu)$ chosen by algorithms A1 and A2 in step 1 satisfies:

$$T'_\nu = M_{\hat{q}(\nu)}$$

and let be

$$\hat{q} \in \bigcap_{n=1}^{\infty} \overline{\{\bigcup_{k \geq n} \hat{q}(k)\}}$$

Then, there exists a subsequence (denoted by ν'), such that:

$$\exists \; n_0 \;/\; \forall \; \nu' \geq n_0 \,, \quad T'_{\nu'} = M_{\hat{q}}$$

since

$$\hat{q}(\nu) \in \hat{Q}(x^{\nu})$$

we have

$$T(x^{\nu}) = M_{\hat{q}}\, x^{\nu} + a_{\hat{q}}$$

and when $\nu \to \infty$, we obtain

$$T\,\bar{x} = M_{\hat{q}}\,\bar{x} + a_{\hat{q}}$$

so $\hat{q} \in \bar{Q}$, and that implies:

$$\bar{x} = (I - M_{\hat{q}})^{-1} a_{\hat{q}}$$

let $\bar{\nu}$ be an index such that $\hat{q}(\bar{\nu}) = \hat{q}$; then in algorithms A1 and A2 with p=1, we have:

$$y^{\bar{\nu},1} = x^{\bar{\nu}} + (I - M_{\hat{q}})^{-1}\,(Tx^{\bar{\nu}} - x^{\bar{\nu}})$$

$$y^{\bar{\nu},1} = x^{\bar{\nu}} + (I - M_{\hat{q}})^{-1}\,(a_{\hat{q}} + (M_{\hat{q}} - I)\,x^{\bar{\nu}} = (I - M_{\hat{q}})^{-1}\,a_{\hat{q}}$$

and then

$$y^{\bar{\nu},1} = \bar{x}$$

in consequence $V(y^{\bar{\nu},1}) = 0$ and in step $\bar{\nu}+1$, algorithm finishes.

□

Remark: In the case T polyhedric condition (9) can be replaced by the simple condition

$$V(y) < V(x)$$

The property of convergence remains valid.

5. NEGATIVE COUNTEREXAMPLE

5.1 Example with at most a Geometric Rate of Convergence of Order 1/3.

• Definition of operator T:

Let be $\frac{1}{3} < \hat{\gamma} < \gamma < 1$, such that

$$\frac{1}{3}(1-\hat{\gamma}) \leq \frac{(1-\gamma)^2}{2} \frac{1}{1-\gamma+4(\gamma-\hat{\gamma})} \tag{22}$$

We define function T in the interval

$$I_b = \left(\frac{\hat{\gamma}}{\gamma} \cdot \frac{1-\gamma}{1-\gamma+4(\gamma-\hat{\gamma})}, 1\right]$$

in the following way:

For $\beta_1 = \frac{\hat{\gamma}}{\gamma}$, $\beta_2 = \frac{1-\gamma}{1-\gamma+4(\gamma-\hat{\gamma})}$ and $\beta = \beta_1 \beta_2 = \frac{\hat{\gamma}}{\gamma} \frac{1-\gamma}{1-\gamma+4(\gamma-\hat{\gamma})}$, we take:

$$T(x) = \begin{cases} \frac{1+3\gamma}{4}x + \frac{\gamma-1}{4} & \text{if } \beta_2 \leq x \leq 1 \\ \hat{\gamma}\beta_2 & \text{if } \beta \leq x < \beta_2 \end{cases}$$

For a general point $x>0$, we define $q(x)=[\frac{\ln x}{\ln \beta}]$ and

$$T(x) = \begin{cases} \frac{1+3\gamma}{4}x + \frac{\gamma-1}{4}\beta^q & \text{if } \beta_2\beta^q \leq x \leq \beta^q \\ \gamma\beta^{q+1} & \text{if } \beta^{q+1} \leq x < \beta_2\beta^q \end{cases}$$

If $x<0$ we define $T(x)=-T(-x)$. Function T has then the shape shown in Figure 1.

• Effect of algorithm A1 on function T:

We shall show that for the special above defined function T, algorithm A1 never get a superlinear rate of convergence but merely a geometric convergence of rate 1/3.

In fact, in a first place we shall proof that algorithm A1 leaves loop 2-3-4-2 always with $p=1$

* If $x^\nu \in [\beta_2\beta^q, \beta^q]$, then:

$$T' = \frac{1+3\gamma}{4},$$

$$v^{\nu,1} = \frac{4}{3(1-\gamma)}\left(T(x^\nu) - x^\nu \right)$$

and

$$y^{\nu,1} = x^\nu + v^{\nu,1} = x^\nu + \frac{4}{3(1-\gamma)}\left(\frac{1+3\gamma}{4}x^\nu + \frac{\gamma-1}{4}\beta^q - x^\nu\right) = -\tfrac{1}{3}\beta^q$$

By definition

$$\hat{\gamma}x \le Tx \le \gamma x$$

$$(1-\gamma)x \le x - Tx \le (1-\hat{\gamma})x \quad \forall\, x \ge 0$$

Then

$$(1-\gamma)^2 x^2 \le \|x - Tx\|^2 \le (1-\hat{\gamma})^2 x^2 \tag{23}$$

In the same way it can be proved that (23) is valid for $x < 0$

$$V(y^{\nu,1}) = \|y^{\nu,1} - Ty^{\nu,1}\|^2 \le (1-\hat{\gamma})^2 \|y^{\nu,1}\|^2 = \tfrac{1}{9}(1-\hat{\gamma})^2\beta^{2q}$$

$$V(x^\nu) \ge (1-\gamma)^2\|x^\nu\|^2 \ge (1-\gamma)^2\beta_2^2\beta^{2q}$$

Then $y^{\nu,1}$ satisfies the test

$$V(y^{\nu,1}) \le \tfrac{1}{2}V(x^\nu) \le \frac{1+\hat{\beta}_\nu}{2}V(x^\nu)$$

because

$$V(y^{\nu,1}) \le \tfrac{1}{9}(1-\hat{\gamma})^2\beta^{2q} \le \frac{(1-\gamma)^2}{2}\ \frac{(1-\gamma)^2}{(1-\gamma+4(\gamma-\gamma'))^2}\,\beta^{2q} \le \frac{V(x^\nu)}{2}$$

by virtue of (22).

Moreover,

$$|y^{\nu,1}| = \frac{\beta^q}{3} \ge \frac{|x^\nu|}{3} \tag{24}$$

* If $x^\nu \in [\beta^{q+1}, \beta_2\beta^q]$, then $T' = 0$,

$$v^{\nu,1} = T(x^\nu) - x^\nu$$

and

$$y^{\nu,1} = T(x^\nu)$$

and then obviously it is verified

$$V(y^{\nu,1}) < \frac{1+\hat{\beta}_\nu}{2} V(x^\nu)$$

Moreover,

$$\|y^{\nu,1}\| \geq \hat{\gamma} \| x^\nu \| \geq \frac{1}{3} \| x^\nu \| \tag{25}$$

In consecuence, it is always verified that $x^{\nu+1} = y^{\nu,1}$

The operator M of the abstract algorithm Aa (that comprises algorithms A1 and A2) verifies, by virtue of (24) and (25):

$$| M(x^\nu) | \geq \frac{1}{3} |x| \tag{26}$$

and in consecuence the convergence rate is never superlinear independently of the chosen starting point x_0.

□

6. A COMPUTATIONAL EXAMPLE

We deal here with a discrete version of a differential games problem, where to find the value function u it is necessary to solve the fixed point problem:

$$u = Tu$$

where

$$Tu = \min_\sigma \max_\alpha (\gamma A^{\alpha,\sigma} u + b^{\alpha,\sigma}) \tag{27}$$

with

$$0 \leq \gamma \leq 1$$
$$\alpha \in \mathcal{A}, \text{card}(\mathcal{A}) = m_1$$
$$\sigma \in \mathcal{S}, \text{card}(\mathcal{S}) = m_2$$

$A^{\alpha,\sigma}$ nxn matrix verifying;

$$A_{i,j}^{\alpha,\sigma} \geq 0$$

$$\sum_j A_{i,j}^{\alpha,\sigma} = 1$$

$$b^{\alpha,\sigma} \in \Re^n$$

It can be easily proved that:

$$T' = \gamma\, A^{\bar{\alpha},\bar{\sigma}}$$

where $\bar{\alpha}$, $\bar{\sigma}$ are the parameters which realize the min-max in (27)

In the examples solved data $A^{\alpha,\sigma}$, $b^{\alpha,\sigma}$ have been generated ramdomly. In the following tables are shown the numerical results and the computational times.

Example 1: $n = 20$, $m_1 = 5$, $m_2 = 5$, $\gamma = 0.999999999$

iterations	V
1	$0.1314 \quad 10^{20}$
2	$0.7248 \quad 10^{4}$
3	0.1987
4	$0.2255 \quad 10^{-2}$
5	$0.7632 \quad 10^{-15}$

Computational time: 11" (PC IBM/AT)

Example 2: $n = 10$, $m_1 = 5$, $m_2 = 5$, $\gamma = 0.999999999$

iterations	V
1	$0.6727 \quad 10^{18}$
2	$0.5173 \quad 10^{3}$
3	0.1512
4	$0,2631 \quad 10^{-2}$
5	$0.1363 \quad 10^{-5}$
6	$0.1110 \quad 10^{-15}$

Computational time: 3" (PC IBM/AT)

CONCLUSIONS

The principal results obtained in this paper are the following:

- For a general nonlinear fixed point problem the obtained algorithms are of Newton's type and they converge from every starting point.
- In the case where the operator $T \in C^1 \bigcap H^{2,\infty}$ quadratic convergence is proved.
- In the case where T is polyhedric convergence in a finite number of steps is obtained.
- Examples are given, showing that it is not posible to define a convergent algorithm of pure Newton's type, and that it is not posible in general to obtain an algorithm with superlinear convergence.

APPENDIX

The methods here proposed enable us to modify the Howard's method (originally created to solve optimal control problems), in order to make it applicable to differential games problems, because for these problems it is in general not convergent.

To show this limitation of Howard's method, we consider an example where the value function of the game is given by the solution of the following fixed point problem.

$$x = Tx$$

where

$$Tx = \min_{a \in A} \max_{b \in B} \left(\beta_{ab} x + c_{ab} \right)$$

being $x \in \Re$, $A = \{0, 1\}$, $B = \{0, 1\}$

$$\beta_{ab} = 3/5 \text{ , } c_{ab} = -\ 6/5 \text{ if } ab = 11$$

$$\beta_{ab} = 0 \text{ , } c_{ab} = -\ 6/5 \text{ if } ab = 00$$

$$\beta_{ab} = 3/5 \text{ , } c_{ab} = 6/5 \text{ if } ab = 01$$

$$\beta_{ab} = 3/5 \text{ , } c_{ab} = 0 \text{ if } ab = 10$$

It is obvious that $| Tx - T\hat{x} | \leq 3/5 | x - \hat{x} |$, and so T is a contractive operator.
Then, the problem is to find a solution of

$$x = \min_{a \in A} \max_{b \in B} \left(\phi_{ab}(x) \right)$$

with

$$\phi_{00}(x) = 0$$

$$\phi_{01}(x) = \frac{3}{5}(x+2)$$

$$\phi_{10}(x) = \phi_{11}(x) = \frac{3}{5}(x-2)$$

If we try to apply a naive version of Howard's method , we would obtain:

$$\text{for } x_0 > 2\,, \quad x_1 = x_0 + (1 - T'(x_0))\,(Tx_0 - x_0) = -3$$

$$\text{for } x_0 < -2\,, \quad x_1 = x_0 + (1 - T'(x_0))\,(Tx_0 - x_0) = 3\,.$$

Then, this procedure would generate a non-convergent sequence.

Algorithms A1 and A2 avoid this phenomenon; in fact the point given by Newton's method (Howard's methods in this case) for $|x| > 2$ is not chosen because the test:

$$V(y^{\nu,1}) < \alpha^{\nu,1}\, V(x^{\nu})$$

is not verified, and algorithms A1 and A2 choose others suitable points and finish in a finite number of steps because T is polyhedrical (see theorem 5.2).

Acknowledgments

The authors would like to thank:

- E. Mancinelli for her skillful preparation of the computer program to run the numerical examples and her almost flawless revision of the manuscript.
- S. Maniscalco for her careful typing of the manuscript.
- CONICET for support given to this work through the grants: PID N° 3-090900/88 and PID N° 3-091000/88.
- The authorities of INRIA for the support given through the Cooperation Project INRIA-Instituto de Matemática Beppo Levi.

REFERENCES

[1] Belbas, S. A., Mayergoyz, I. D.: Applications of fixed-point methods to discrete variational and quasi-variational inequalities. *Numerische Mathematik*, Vol. 51, pp. 631-654, 1987.

[2] Brezinski C.: *Accélération de la Convergence en Analyse Numérique* (Springer-Verlag, New York, 1977).

[3] Clarke, F. H.: *Optimization and Nonsmooth Analysis* (Wiley , New York, 1983).

[4] El Tarazi, M. N.: On a monotony-preserving accelerator process for the successive approximations method. *IMA Journal of Numerical Analysis*. Vol. 6, pp. 439-446, 1986.

[5] Glowinski R.: *Numerical Methods for Nonlinear Variational Problems* (Springer-Verlag, New York, 1984).

[6] Glowinski R., Lions J. L., Trémolières R.: *Numerical Analysis of Variational Inequalities* (North Holland, Amsterdam, 1981).

[7] González R., Rofman E.: On deterministic control problems: an approximation procedure for the optimal cost. Part I and II. *SIAM Journal on Control and Optimization*, 23, pp. 242-285, 1985.

[8] González, R., Sagastizábal, C.: On subsolutions and supersolutions of discrete Hamilton-Jacobi-Bellman equations: An accelerated algorithm to solve optimal control problems, *Rapport de Recherche* N° 1014, INRIA, 1989.

[9] González, R., Sagastizabál, C.: Un algorithme pour la résolution rapide d'équations discrètes de Hamilton-Jacobi-Bellman. *Comptes Rendus Acad. Sc. Paris*, Serie I, Tome 311, pag. 45-50, 1990.

[10] González, R., Tidball, M.: Fast solution of discrete Isascs's inequalities, *Rapport de Recherche* N° 1167, INRIA, 1990.

[11] Lemaréchal, C.: Extensions diverses des méthodes de gradient et applications, Thèse d'Etat, UER Mathématiques de la Décision, University of Paris IX (Paris, 1980).

[12] Lemaréchal C.: Basic Theory in Nondifferentiable Optimization, *Rapport de Recherche* N° 181, INRIA, 1982.

[13] Marcotte P.: A new algorithm for solving variational inequalities with application to the traffic assignment problem. *Mathematical Programming*, Vol. 33, pp. 339-351, 1985.

[14] Mosco U.: An introduction to the approximate solution of variational inequalities. In *Constructive Aspects of Functional Analysis*, Vol II, pp. 497-682, (Edizioni Cremonese, Roma, 1971).

[15] Ortega, J. M., Rheinboldt, W. C.: *Iterative solution of nonlinear equations in several variables* (Academic Press, New York, 1970).

[16] Polak E.: *Computational Methods in Optimization. A Unified Approach* (Academic Press, New York, 1970).

[17] Ross S. M.: *Applied Probability Models with Optimization Applications* (Holden-Day, San Francisco, 1970).

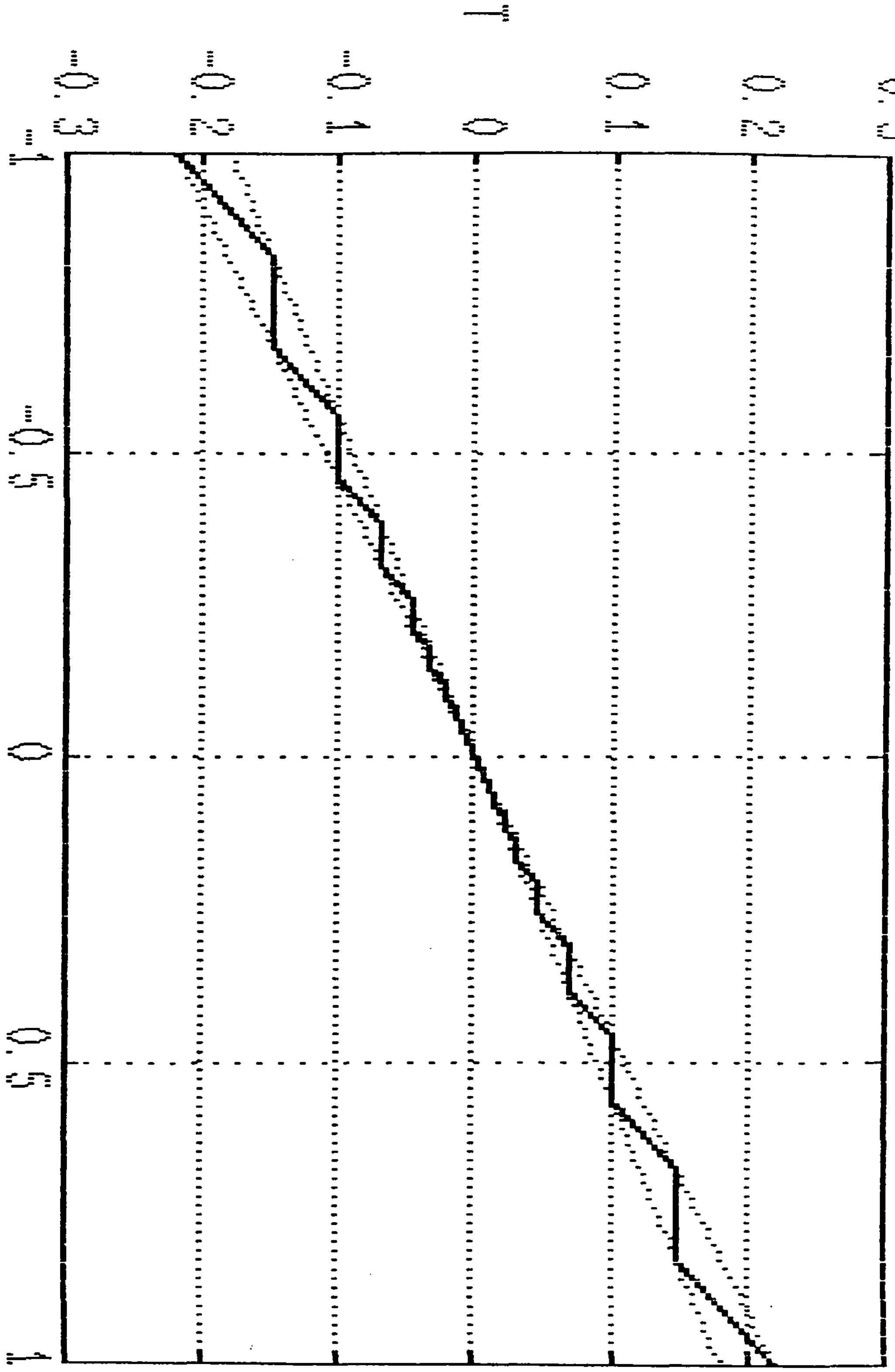

FIGURE 1

Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique

ISSN 0249-6399